Reactive-Ion Plasma Etching Process

Introduction:

Multi-layer perceptron's are universal approximators and they can learn any non-linear relationship. MLP networks are feed forward networks i.e the signal propagates through the network in a forward direction, on layer by layer basis. Typically, these networks are trained in a supervised fashion using error back-propagation algorithm based on the error-correction learning rule. Error back-propagation consists of two passes such as forward and backward passes.

Below are the common steps fallowed to understand the data and models development for both the datasets.

1.Analyze data to understand the variables and relationships

   a. Understand relationship between the independent and dependent variables
   b. Analyze the density of independent and dependent variables variables. Perform normalization if necessary

2.Develop basic model to understand the learning patterns of the MLP and performed further assessment with different model variations and improve the model with various parameters such as additional hidden layers, Regularizations, non-linearity activation functions, K-fold cross validations.

   a. Observe the model performance with more hidden layers
   b. Apply Regularization on the model and observe evaluation parameters such as RMSE, Accuracy
   c. Perform cross validation (K=10, LeaveOneOut) and observe model performance

3.Use Skopt optimizer to find better hyper parameters such as number of neurons for each layer and drop-out rate

4.Use cross validation and hyper parameter tuning to determine best hyper parameters and metrices.

5.Test the final model

Reactive-Ion Plasma Etching Process

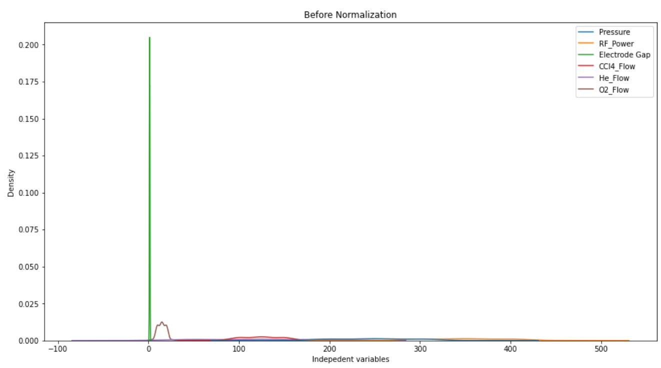
In this problem, there are 53 data points and 11 variables. The objective is to develop a Multi-Layer Perceptron neural network that indicates the behavior of an ion-assisted plasma etching semiconductor fabrication process. Below is the summary of the dataset.

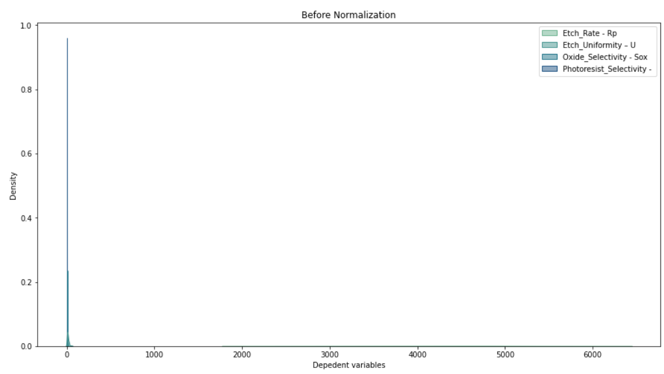
	Pressure	RF_Power	Electrode Gap	CCl4_Flow	He_Flow	O2_Flow	Etch_Rate - Rp	Etch_Uniformity – U	Oxide_Selectivity - Sox	Photoresist_Selectivity -
count	53.0	53.0	53.0	53.0	53.0	53.0	53.0	53.0	53.0	53.0
mean	250.0	350.0	1.5	125.0	124.1	15.0	4204.3	11.6	7.0	2.2
std	45.6	45.6	0.3	22.9	62.2	4.6	689.4	10.6	2.2	0.5
min	131.0	231.0	0.8	64.0	0.0	3.0	2704.0	0.5	2.7	1.3
0.25	200.0	300.0	1.2	100.0	50.0	10.0	3684.0	3.9	5.8	2.0
0.5	250.0	350.0	1.5	125.0	125.0	15.0	4390.0	8.3	6.4	2.1
0.75	300.0	400.0	1.8	150.0	200.0	20.0	4703.0	15.1	7.7	2.3
max	369.0	469.0	2.2	184.0	200.0	27.0	5515.0	55.2	15.2	4.2

There are some independent and dependent variables are correlated. For example, Oxide_Selectivity – Sox is negatively correlated with He_Flow dependent variable. Below correlation matrix shows the relationship between independent and dependent variables.

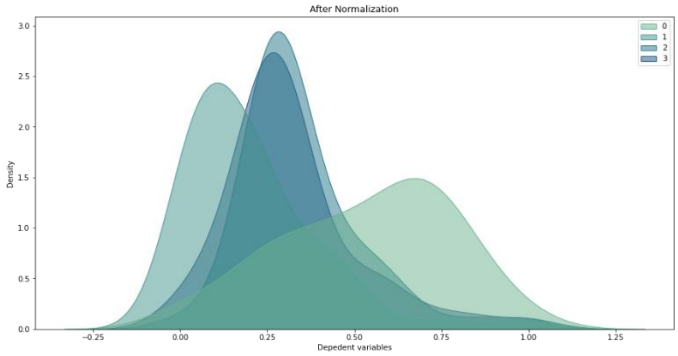
Independent and dependent variables are not normally distributed in the dataset. So, it is a good idea to normalize the data for further analysis.

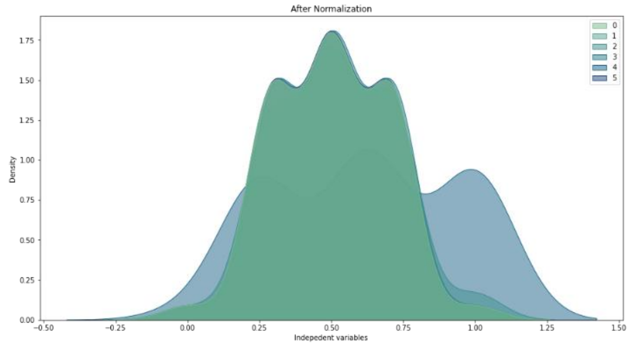
Before Normalization:





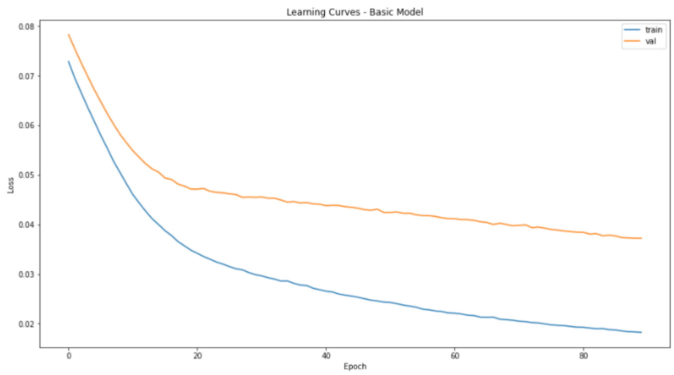
After Normalization:

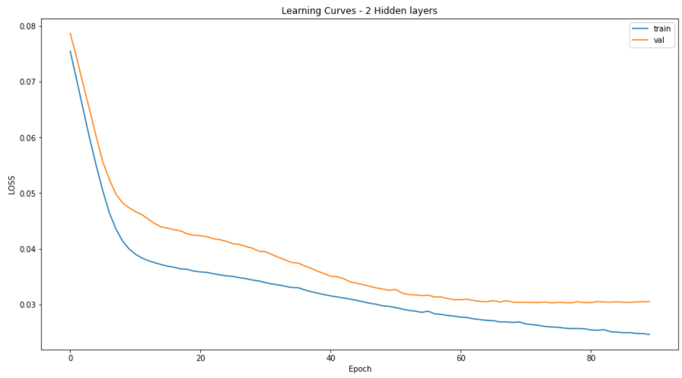




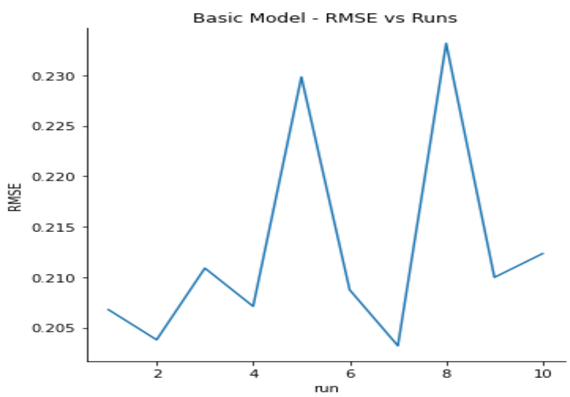
Basic Model:

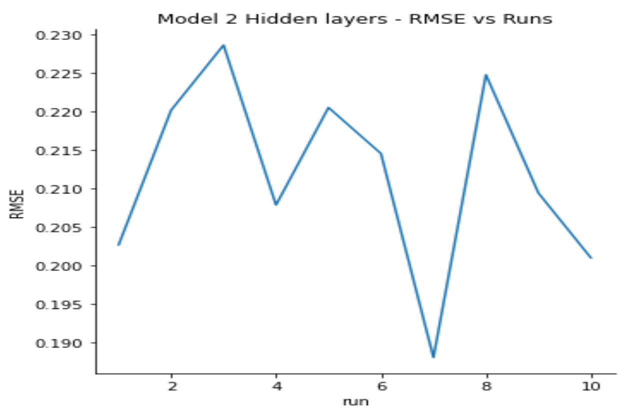
Basic Model consists of input layer with 7 neurons, one hidden layer with 9 neurons and output layer consists of 4 neurons. Given dataset is divided into 80% training data and 20% testing dataset. "ReLu" is considered as activation function for input, hidden layers and "`sigmoid`" is used as output layer activation function. Basic model is converging around 20-25 Epochs.





Basic model and model with 2 hidden layers RMSE values are varying between 19% and 23 % so adding more layers may not yielded any improvement in the model. Below graphs shows the RMSE values at different runs for both models.



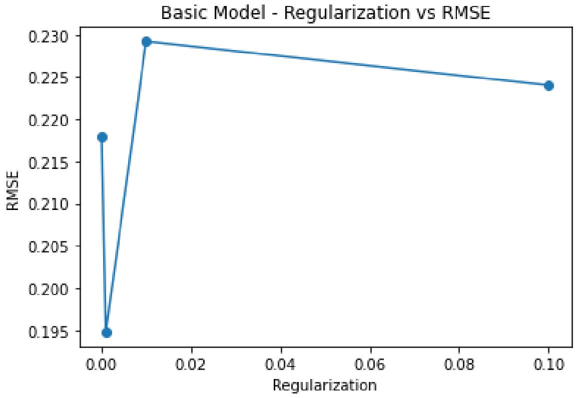


Regularization:

Basic model is considered for regularization because adding more hidden layers is not improving the model performance. This is observed in earlier analysis. Following parameters are used for the model.

- Layers: Input layers, one hidden layer and output layer
- No.of neurons: input layer=7,hidden layer=9 and output layer=4
- Activation function: input, hidden layers="ReLu" and output layer="Sigmoid"
- Regularization: L1

Below plot indicated Regularization didn't offer much help to improve the model performance. RMSE value is varied between 19% to 23 % at different L1 values.

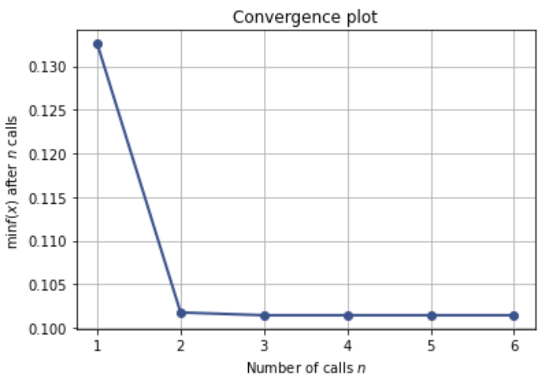


Hyper Parameter Tuning:

Hyper parameter tuning is used to find out optimal neurons to the input and hidden layers. Below are the configuration steps used to find out best model and optimal values. "Skopt" is used for hyper parameter tuning.

- SPACE is configured with input layer neuron range 4-1200 and hidden layer with 9-600
- This model has input,one hidden and output layers
- No of Epochs are 90
- Activation function: input, hidden layers="ReLu" and output layer="Sigmoid"
- No of calls is 6

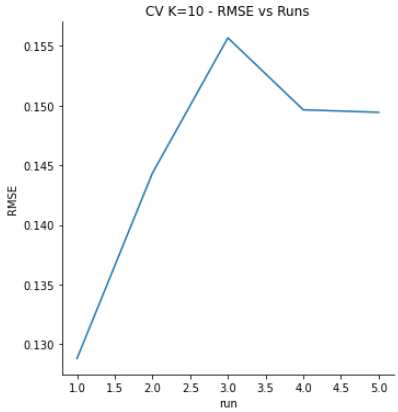
Best model RMSE value is 10% with `1030` input neurons and 510 hidden layer neurons and it is achieved at iteration 2. Below graph shows RMSE values for each iteration of hyper parameter tuning and optimal parameters are changing for each run so decided to use cross validation within the skopt model. Even though this is resulted less RMSE. This model is likely memorizing it so it is a good idea to check with cross validation.



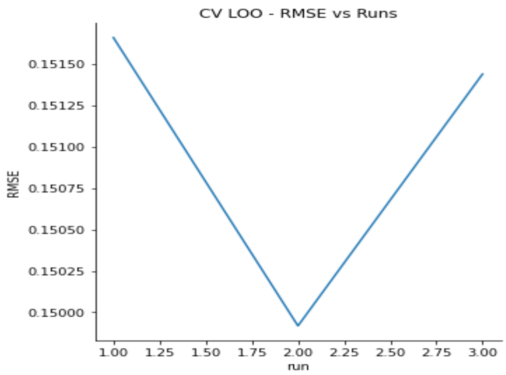
Cross validation:

Given dataset is sparse so using of cross validation may give us the better model performance. K-10 and leave one out option are considered to observe the model performance with Basic model parameter configurations are used as mentioned above.

K=10-fold cross validation improved the performance of the model. RMSE value is approx..15% with std of +/- 2%. Below graph shows the plot of cross validation model for multiple runs.



Leave One Out cross validation didn't show much improvement in model performance compared to K=10 cross validation. Below plot shows RMSE values for 3 runs.



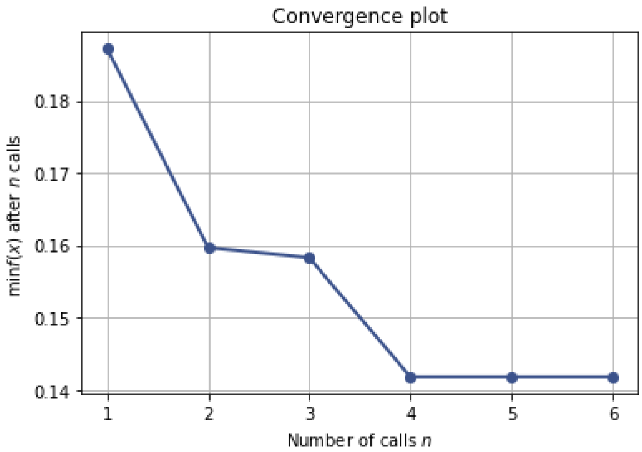
Hyper parameter tuning & Cross validation:

Developed another model to find the best performance and optimal parameters with cross validation K=10 and Leave one out. Below are the parameters used for this model.

- SPACE is configured with input layer neuron range 4-1200 and hidden layer with 9-600,drop-out range is 0 to 0.3
- This model has input,one hidden and output layers
- No of Epochs are 90
- Activation function: input, hidden layers="ReLu" and output layer="Sigmoid"
- No of calls is 6

➢ K=10 cross validation & Leave one out cross validation

This model resulted RMSE 14% with 50 input neurons,59 hidden neurons and 0.14 is the drop-out rate for K=10. Below is the plot of RMSE for different calls.



Leave one out and skopt optimization model generated 12% RMSE and this is best model so far with 85 input neurons,47 hidden layer neurons and 0.11 is the drop rate.

Final Model and conclusion:

Developed final model with cross validation(LOOCV) and with optimized neurons obtained from the hyper parameter tuning and this model consistently resulted 11% RMSE with std +/- 1%..Below plot is for final model for 3 runs with RMSE values and this is the best model.